

Introdução à Ciência de Dados e Métodos Quantitativos

Estudo de Caso

Você é parte de um time de ciência de dados em um grande banco com sede em diversos países. O gerente de relacionamentos do banco ouviu sobre “métodos quantitativos” em um conferência e começou a pensar se essa ferramenta poderia ajuda-lo em suas tarefas. Uma vez que o time de relacionamentos não dispõe de analistas ciêntistas de dados eles te convidaram para uma reunião para entender as possibilidades. Eles pediram para você “analisar os dados” e te enviaram de antemão dois conjuntos de dados.

O primeiro conjunto, intitulado `Retail_data.csv` contém informações sobre clientes escolhidos aleatoriamente que receberam ou não um empréstimo ao longo dos últimos anos, além de dados adicionais sobre os clientes como a quantidade de anos que o cliente está com o banco, estado marital, nível de educação, status de empregabilidade, renda mensal e balanço total na conta.

O segundo conjunto de dados, intitulado `Potential_Customers.csv` contém os mesmos dados sobre potenciais clientes exceto a informação sobre o empréstimo. O gerente de relacionamentos gostaria de saber quais potenciais clientes ele deveria abordar para oferecer um empréstimo.

Sua tarefa

1 – Prepare um conjunto de perguntas com base nos dados para iniciar uma conversa durante a reunião. (25% da nota final)

2 – Realize uma análise exploratória dos dados, indicando padrões interessantes para o cliente de relacionamentos. (25% da nota final)

3 – Crie um modelo para prever quais potenciais clientes deveriam ser abordados. (40% da nota final)

4 – Valide seu(s) modelo(s) utilizando o conceito de matriz de confusão. (10% da nota final).